Weak uniform structures on probability distributions

Jan Pachl

Fields Institute*

Toronto, Ontario, Canada

November 21, 2010 (version 2)

Abstract

In dealing with asymptotic approximation of possibly divergent nets of probability distributions, we are led to study uniform structures on the set of distributions. This paper identifies a class of such uniform structures that may be considered to be reasonable generalizations of the weak topology. It is shown that all structures in the class yield the same notion of asymptotic approximation for sequences (but not for general nets) of probability distributions.

1 Introduction

Traditionally limit theorems in probability theory are formulated in terms of the *weak topology* on probability distributions, also known as the topology of convergence in distribution. When a sequence of probability distributions converges in the weak topology, its limit serves as a *weak approximation* of the distributions in the sequence.

In more general limit theorems, a sequence of probability distributions, not necessarily convergent, is asymptotically approximated by another sequence. That leads to a natural question: Are there essentially different "reasonable" notions of asymptotic approximation that all reduce to the weak approximation in the special case of convergent sequences?

To deal with approximation for possibly divergent sequences we need more than a mere topology. Asymptotic approximation is conveniently formulated in terms of uniform structures (uniformities) on the space of probability distributions. The question above then becomes: What "reasonable" uniform structures on the space of probability distributions are compatible with the weak topology?

Uniform structures on probability distributions on a metric space were investigated by Dudley [6][8, 11.7] and D'Aristotile, Diaconis and Freedman [3]. Davydov and Rotar [4] showed that the previously neglected uniformity defined by bounded uniformly continuous functions is a reasonable uniformity in the above mentioned sense.

---

*Paper written while the author was a visitor at the Fields Institute.

Uniform structures on probability distributions are also implicit in the rich literature on probability metrics, surveyed by Dudley [7], Gibbs and Su [10], Rachev [16] and Zolotarev [19]; however, so far the focus in that area has been on quantitative results for convergent sequences, not on general uniform structures or even the uniform structures defined by probability metrics.

In this paper I offer further evidence that the uniformity proposed by Davydov and Rotar is an appropriate analog of the weak topology, at least when we deal with sequences (rather than general nets) of probability distributions. To formalize the notion of "reasonable" uniform structures, I formulate two natural properties that they should satisfy. Although there is usually more than one uniformity satisfying those properties, and therefore more than one notion of a "weak" asymptotic approximation for *nets* of probability distributions, they all agree on *sequences*.

Version 1 of the paper was dated July 20, 2010. Version 2 incorporates a suggestion from Ramon van Handel, which considerably simplifies and improves the results in section 3.

2 Preliminaries

Uniform structures (uniformities) and uniform spaces may be defined in several equivalent ways. For the purposes of this paper, the most suitable definition is the one based on pseudometrics [11, Ch.15], although definitions in terms of entourages [1] or uniform covers [12] would do as well. Let $\mathcal{U}$ be a uniform structure on a set A. A net $\{a_\gamma\}_\gamma$ in A is an *asymptotic $\mathcal{U}$-approximation* of a net $\{b_\gamma\}_\gamma$ in A iff $\lim_\gamma \Delta(a_\gamma, b_\gamma) = 0$ for every $\mathcal{U}$-uniformly continuous pseudometric Δ. Here a *net* is a family indexed by a directed partially ordered set. A *sequence* is a family indexed by the totally ordered set $\{0, 1, 2, \dots\}$.

The set of real numbers is denoted $\mathbb{R}$. When S a metric space, $\mathsf{C_b}(S)$ is the space of bounded real-valued continuous functions on S, and $\mathsf{U_b}(S)$ is the space of bounded real-valued uniformly continuous functions on S. The sup norm of a function $f \in \mathsf{C_b}(S)$ is $\|f\|_S := \sup_{x \in S} |f(x)|$.

The *Borel σ-algebra* $\mathsf{Bo}(S)$ is the smallest σ-algebra of subsets of S containing all open sets in S. In this paper, a *measure on S* means a bounded signed measure on $\mathsf{Bo}(S)$. A measure μ on S is *tight* if

$$|\mu|(A) = \sup\{|\mu|(C) \mid C \text{ is compact and } C \subseteq A\}$$

for every $A \in \mathsf{Bo}(S)$. Tight measures are also known as Radon or regular measures. The space of tight measures on S is denoted $\mathsf{M_t}(S)$, and

$$\mathsf{P_t}(S) := \{\mu \in \mathsf{M_t}(S) \mid \mu \geq 0 \text{ and } \mu(S) = 1\}$$

is the space of tight probability measures on S. The *point mass at $x \in S$* is denoted $\partial_S(x)$; that defines a mapping $\partial_S \colon S \to \mathsf{P_t}(S)$.

The vector space duality $\langle \mathsf{M_t}(S), \mathsf{C_b}(S) \rangle$ is defined by integration:

$$\langle \mu, f \rangle := \int f \, \mathsf{d}\mu \quad \text{for } \mu \in \mathsf{M_t}(S), f \in \mathsf{C_b}(S).$$

Since every $\mu \in \mathsf{M_t}(S)$ is uniquely determined by the values $\langle \mu, f \rangle$, $f \in \mathsf{U_b}(S)$, integration defines also a vector space duality $\langle \mathsf{M_t}(S), \mathsf{U_b}(S) \rangle$. When X is a real-valued random variable, $\mathsf{E}(X)$ denotes the expected value of X. If $f \in \mathsf{C_b}(S)$ and X is an S-valued random variable with distribution $\mu \in \mathsf{P_t}(S)$ then $\langle \mu, f \rangle = \mathsf{E}(f(X))$.

When S and S' are metric spaces and $\varphi \colon S \to S'$ is a continuous mapping, the linear mapping $\mathsf{M_t}(\varphi) \colon \mathsf{M_t}(S) \to \mathsf{M_t}(S')$ is defined by $\langle \mathsf{M_t}(\varphi)(\mu), g \rangle := \langle \mu, g \circ \varphi \rangle$ for $\mu \in \mathsf{M_t}(S)$ and $g \in \mathsf{C_b}(S')$. Clearly $\mathsf{M_t}(\varphi)$ maps $\mathsf{P_t}(S)$ into $\mathsf{P_t}(S')$.

When d is the metric of S and $\mu \in \mathsf{M_t}(S)$, write

$$\mathsf{BLip_b}(d) := \{ f \colon S \to \mathbb{R} \mid \|f\|_S \leq 1 \ \text{ and } \ |f(x) - f(y)| \leq d(x,y) \ \text{ for all } \ x,y \in S \}$$
$$\|\mu\|^* := \sup \{ \langle \mu, f \rangle \mid f \in \mathsf{BLip_b}(d) \} .$$

By the following well-known lemma, $\|\cdot\|^*$ is a norm on the space $\mathsf{M_t}(S)$. This norm is equivalent to the norm $\|\cdot\|_{\mathsf{BL}}^*$ studied by Dudley [5][6].

Lemma 1 *Let S be a metric space with metric d. The space $\bigcup_{n=1}^{\infty} n\mathsf{BLip_b}(d)$ is $\|\cdot\|_S$-dense in the space $\mathsf{U_b}(S)$.*

   Proof. Take any $f \in \mathsf{U_b}(S)$, $\varepsilon > 0$. There is $\theta > 0$ such that if $x, y \in S$ and $d(x,y) < \theta$ then $|f(x) - f(y)| < \varepsilon$. Choose an integer $n \geq \max(\|f\|_S + \varepsilon, 2\|f\|_S/\theta)$ and define

$$g(y) := \sup_{x \in S} \, (f(x) - \varepsilon - n\, d(x,y)) \ \text{ for } \ y \in S.$$

Then $g \in n\mathsf{BLip_b}(d)$ and $f - \varepsilon \leq g \leq f$. $\qquad\qquad\square$

In addition to the $\|\cdot\|^*$ topology, consider also two weak topologies on $\mathsf{M_t}(S)$:

- The $\mathsf{C_b}(S)$-weak topology from the duality $\langle \mathsf{M_t}(S), \mathsf{C_b}(S) \rangle$. That is, the topology of simple convergence on the elements of $\mathsf{C_b}(S)$.

- The $\mathsf{U_b}(S)$-weak topology from the duality $\langle \mathsf{M_t}(S), \mathsf{U_b}(S) \rangle$. That is, the topology of simple convergence on the elements of $\mathsf{U_b}(S)$.

The restriction of the $\mathsf{C_b}(S)$-weak topology to $\mathsf{P_t}(S)$ is called simply the *weak topology* (or the topology of convergence in distribution) in probability theory.

Lemma 2 *Let S be a metric space. The $\mathsf{C_b}(S)$-weak topology, the $\mathsf{U_b}(S)$-weak topology and the $\|\cdot\|^*$ topology coincide on $\mathsf{P_t}(S)$.*

   Proof. LeCam [13, Lem.5] proves that the two weak topologies coincide on $\mathsf{P_t}(S)$. The equivalence with the $\|\cdot\|^*$ topology is proved by LeCam [14], and by Dudley [5, Th.18]. $\qquad\square$

The proof of Corollary 11 in section 3 relies on the following theorem, which generalizes the Schur property [9, 5.19] of convergent sequences in ℓ_1.

Theorem 3 *Let S be a metric space. The $\mathsf{U}_\mathsf{b}(S)$-weak topology and the $\|\cdot\|^*$ topology on $\mathsf{M}_\mathsf{t}(S)$ have the same compact sets, and therefore the same convergent sequences.*

Theorem 3 appeared in [15]. A simpler proof was found by Cooper and Schachermayer [2]. Variants of the theorem were proved by van Handel [18, B.1] (for measures on $\mathbb{R}^n$) and Davydov and Rotar [4, Th. 4].

Theorem 4 *Let S be a metric space and let $\mu_j, \nu_j \in \mathsf{P}_\mathsf{t}(S)$ for $j = 0, 1, \dots$. We have*

$$\lim_j \|\mu_j - \nu_j\|^* = 0$$

if and only if there exist S-valued random variables X_j, Y_j such that the distribution of X_j is μ_j, the distribution of Y_j is ν_j, and $\lim_j d(X_j, Y_j) = 0$ almost surely.

Theorem 4 is proved by Dudley [8, 11.7.1]. (In [8], β denotes the metric of the norm $\|\cdot\|^*_{\mathsf{BL}}$, which is equivalent to $\|\cdot\|^*$.)

When E is a locally convex vector space, the topology of E is defined by the family of continuous seminorms [17, II.4]. The *additive uniformity* on any subset of E is the uniformity induced by the pseudometrics of the form $(x, y) \mapsto s(x - y)$, $x, y \in E$, where s is a continuous seminorm on E. Each of the three topologies on $\mathsf{M}_\mathsf{t}(S)$ defined above makes $\mathsf{M}_\mathsf{t}(S)$ into a locally convex space. The corresponding additive uniformities on subsets of $\mathsf{M}_\mathsf{t}(S)$ (in particular, on $\mathsf{P}_\mathsf{t}(S)$) will be referred to as the $\|\cdot\|^*$ uniformity, the $\mathsf{C}_\mathsf{b}(S)$-weak uniformity and the $\mathsf{U}_\mathsf{b}(S)$-weak uniformity.

Lemma 5 *Let S be a metric space. On $\mathsf{P}_\mathsf{t}(S)$, the $\|\cdot\|^*$ uniformity is finer than the $\mathsf{U}_\mathsf{b}(S)$-weak uniformity.*

Proof. Follows from Lemma 1. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3 Uniform structures on $\mathsf{P}_\mathsf{t}(S)$

In this section I describe a class of uniform structures on $\mathsf{P}_\mathsf{t}(S)$ that appear to be reasonable candidates for extending to divergent nets the notion of weak approximation on convergent nets. Of course, it is highly subjective and context-dependent what uniformities should be considered "reasonable" for this purpose. I replace that subjective notion by the two properties in Definition 7.

Lemma 6 *Let $I_n := [-n, n] \subseteq \mathbb{R}$ with the standard metric. There is a unique uniformity on $\mathsf{P}_\mathsf{t}(I_n)$ compatible with the $\mathsf{C}_\mathsf{b}(I_n)$-weak topology. This unique uniformity coincides with the $\|\cdot\|^*$ uniformity, the $\mathsf{C}_\mathsf{b}(I_n)$-weak uniformity and the $\mathsf{U}_\mathsf{b}(I_n)$-weak uniformity on $\mathsf{P}_\mathsf{t}(I_n)$.*

Proof. Since I_n is compact, so is $\mathsf{P}_\mathsf{t}(I_n)$ with the $\mathsf{C}_\mathsf{b}(I_n)$-weak topology. It follows that there is only one uniformity on $\mathsf{P}_\mathsf{t}(I_n)$ compatible with the $\mathsf{C}_\mathsf{b}(I_n)$-weak topology [12, II.24], and by Lemma 2 it coincides with the three uniformities listed. $\qquad\square$

Property (A1) in the next definition and the resulting simplification of the proofs that follow were suggested by R. van Handel.

Definition 7 *Let S be a metric space with metric d. Consider the following properties of a uniform structure $\mathcal{U}$ on $\mathsf{P}_\mathsf{t}(S)$.*

(A1) If X_j and Y_j, $j = 0, 1, \ldots$, are S-valued random variables with distributions $\mu_j \in \mathsf{P}_\mathsf{t}(S)$ and $\nu_j \in \mathsf{P}_\mathsf{t}(S)$ respectively and if $\lim_j d(X_j, Y_j) = 0$ almost surely then the sequence $\{\mu_j\}_j$ is an asymptotic $\mathcal{U}$-approximation for the sequence $\{\nu_j\}_j$.

(A2) For every interval $I_n := [-n, n] \subseteq \mathbb{R}$, $n = 1, 2, \ldots$, if a mapping $\varphi \colon S \to I_n$ is uniformly continuous then so is the mapping $\mathsf{M}_\mathsf{t}(\varphi)$ from $\mathsf{P}_\mathsf{t}(S)$ with $\mathcal{U}$ to $\mathsf{P}_\mathsf{t}(I_n)$ with the unique uniformity in Lemma 6.

Property (A1) states that asymptotic approximation (understood almost surely) for sequences of S-valued random variables implies asymptotic approximation for the corresponding distributions.

(A2) is a functorial property of the assignment $X \mapsto (\mathsf{P}_\mathsf{t}(S), \mathcal{U})$ for uniformly continuous mappings $\varphi \colon S \to I_n$, assuming that on convergent nets in $\mathsf{P}_\mathsf{t}(I_n)$ the asymptotic approximation agrees with the weak approximation.

As a special case of (A1), if $\{x_j\}_j$ and $\{y_j\}_j$ are two sequences of points in S such that $\lim_j d(x_j, y_j) = 0$ then $\{\partial(x_j)\}_j$ is an asymptotic $\mathcal{U}$-approximation for $\{\partial(y_j)\}_j$. Consider the points $x_j := j$ and $y_j = j + 1/j$, $j = 1, 2, \ldots$, in $\mathbb{R}$. There is a function $f \in \mathsf{C}_\mathsf{b}(\mathbb{R})$ such that $f(x_j) - f(y_j) = 1$ for all j; it follows that the $\mathsf{C}_\mathsf{b}(\mathbb{R})$-weak uniformity on $\mathsf{P}_\mathsf{t}(\mathbb{R})$ does not satisfy (A1).

The other two uniformities from section 2 have properties (A1) and (A2):

Theorem 8 *For every metric space S the $\mathsf{U}_\mathsf{b}(S)$-weak uniformity and the $\|\cdot\|^*$ uniformity on $\mathsf{P}_\mathsf{t}(S)$ have properties (A1) and (A2).*

Proof. Let $\mu_j, \nu_j \in \mathsf{P}_\mathsf{t}(X)$, $j = 0, 1, \ldots$, be the distributions of S-valued random variables X_j, Y_j such that $\lim_j d(X_j, Y_j) = 0$ almost surely. Then

$$|\langle \mu_j, f \rangle - \langle \nu_j, f \rangle| = |\mathsf{E}(f(X_j)) - \mathsf{E}(f(Y_j))| \le \mathsf{E}(|f(X_j) - f(Y_j)|)$$
$$\|\mu_j - \nu_j\|^* = \sup_{f \in \mathsf{BLip}_\mathsf{b}(d)} |\langle \mu_j, f \rangle - \langle \nu_j, f \rangle| \le \mathsf{E}(2 \wedge d(X_j, Y_j))$$

and if $\lim_j d(X_j, Y_j) = 0$ almost surely then $\lim_j \|\mu_j - \nu_j\|^* = 0$. Hence the $\|\cdot\|^*$ uniformity has property (A1), and so does the $\mathsf{U}_\mathsf{b}(X)$-weak uniformity by Lemma 5.

(A2) follows from the definition of $\mathsf{M}_\mathsf{t}(\varphi)$ and Lemma 6. $\qquad\square$

Next I prove that every uniformity $\mathcal{U}$ satisfying (A1) and (A2) is between the two uniformities in Theorem 8.

Theorem 9 *Let X be a metric space with metric d, and $\mathcal{U}$ a uniform structure on $\mathsf{P}_\mathsf{t}(X)$.*

1. *If $\mathcal{U}$ has property (A1) then it is coarser than the $\|\cdot\|^*$ uniformity on $\mathsf{P}_\mathsf{t}(X)$.*

2. *If $\mathcal{U}$ has property (A2) then it is finer than the $\mathsf{U}_\mathsf{b}(X)$-weak uniformity on $\mathsf{P}_\mathsf{t}(X)$.*

Proof. To prove part 1, assume that $\mathcal{U}$ is not coarser than the $\|\cdot\|^*$ uniformity on $\mathsf{P}_\mathsf{t}(X)$. That means that there are a $\mathcal{U}$-uniformly continuous pseudometric Δ, $\varepsilon > 0$ and $\mu_j, \nu_j \in \mathsf{P}_\mathsf{t}(X)$ for $j = 0, 1, \ldots$ such that $\lim_j \|\mu_j - \nu_j\|^* = 0$ and $\Delta(\mu_j, \nu_j) \geq \varepsilon$ for all j. By Theorem 4 there are S-valued random variables X_j, Y_j whose distributions are μ_j, ν_j respectively and such that $\lim_j d(X_j, Y_j) = 0$ almost surely. Thus $\mathcal{U}$ does not have property (A1).

2. Assume that $\mathcal{U}$ has property (A2) and take any $f \in \mathsf{U}_\mathsf{b}(X)$. Choose $n \in \{1, 2, \ldots\}$ for which $\|f\|_X \leq n$, so that f maps X into $I_n := [-n, n]$.

Let $g \in \mathsf{U}_\mathsf{b}(I_n)$ be the function $g : x \mapsto x$. By (A2) and Lemma 6 the mapping $\mathsf{M}_\mathsf{t}(f)$ is uniformly continuous from $\mathsf{P}_\mathsf{t}(X)$ with $\mathcal{U}$ to $\mathsf{P}_\mathsf{t}(I_n)$ with the $\mathsf{U}_\mathsf{b}(I_n)$-weak uniformity. Thus the mapping

$$\mu \mapsto \langle \mathsf{M}_\mathsf{t}(f)(\mu), g \rangle = \langle \mu, g \circ f \rangle = \langle \mu, f \rangle$$

from $\mathsf{P}_\mathsf{t}(X)$ with $\mathcal{U}$ to $\mathbb{R}$ is uniformly continuous. That proves that $\mathcal{U}$ is finer than the $\mathsf{U}_\mathsf{b}(X)$-weak uniformity. $\square$

Corollary 10 *Let X be a metric space with metric d, and $\mathcal{U}$ a uniform structure on $\mathsf{P}_\mathsf{t}(X)$. If $\mathcal{U}$ has properties (A1) and (A2) then it is compatible with the $\mathsf{C}_\mathsf{b}(X)$-weak topology on $\mathsf{P}_\mathsf{t}(X)$.*

Proof. Apply Theorem 9 and Lemma 6. $\square$

Corollary 11 *Let S be a metric space and $\mathcal{U}$ a uniform structure on $\mathsf{P}_\mathsf{t}(S)$. If $\mathcal{U}$ has properties (A1) and (A2) then the following statements are equivalent for any two sequences $\{\mu_j\}_j$ and $\{\nu_j\}_j$ in $\mathsf{P}_\mathsf{t}(S)$.*

(i) $\lim_j (\langle \mu_j, f \rangle - \langle \nu_j, f \rangle) = 0$ *for every $f \in \mathsf{U}_\mathsf{b}(S)$.*

(ii) *The sequence $\{\mu_j\}_j$ is an asymptotic $\mathcal{U}$-approximation of the sequence $\{\nu_j\}_j$.*

(iii) $\lim_j \|\mu_j - \nu_j\|^* = 0$.

(iv) *There exist S-valued random variables $X_j, Y_j, j = 0, 1, \ldots,$ whose distributions are μ_j and ν_j and such that $\lim_j d(X_j, Y_j) = 0$ almost surely.*

Proof. The implication (i)⇒(iii) follows from Theorem 3, (iii)⇒(ii)⇒(i) from Theorem 9, and (iii)⇔(iv) from Theorem 4. □

The equivalence of (i) and (iii) in the corollary was derived by Davydov and Rotar [4] from their variant of Theorem 3. By [8, 11.7.1], condition (iii) is also equivalent to each of the following:

- $\lim_j \rho(\mu_j, \nu_j) = 0$, where ρ is the Lévy–Prokhorov metric.

- There exist S-valued random variables X_j, Y_j, $j = 0, 1, \ldots$, whose distributions are μ_j and ν_j and such that $\lim_j d(X_j, Y_j) = 0$ in probability.

4   Concluding remarks

By Corollary 11, all uniformities on $\mathsf{P_t}(S)$ that satisfy (A1) and (A2) yield the same notion of asymptotic approximation for *sequences* of tight probability measures, namely the approximation defined by the $\mathsf{U_b}(S)$-weak and $\|\cdot\|^*$ uniformities.

When we look beyond sequences and deal with general *nets* of probability distributions, there are multiple notions of asymptotic approximation, even for the uniformities with properties (A1) and (A2). For example, the $\|\cdot\|^*$ uniformity on $\mathsf{P_t}(\mathbb{R})$ is strictly finer than the $\mathsf{U_b}(\mathbb{R})$-weak uniformity, and thus there are two nets $\{\mu_\gamma\}_\gamma$ and $\{\nu_\gamma\}_\gamma$ in $\mathsf{P_t}(\mathbb{R})$ such that $\{\nu_\gamma\}_\gamma$ is an asymptotic $\mathsf{U_b}(\mathbb{R})$-weak approximation of $\{\mu_\gamma\}_\gamma$ but not an asymptotic $\|\cdot\|^*$ approximation.

References

[1] Bourbaki, N. *Éléments de mathématique. Topologie générale. Ch. I et II.* Hermann & Cie., Paris, 1940.

[2] Cooper, J. B., and Schachermayer, W. *Uniform measures and co-Saks spaces*. Functional analysis, holomorphy, and approximation theory (Rio de Janeiro, 1978), Lecture Notes in Math., Vol. 843. Springer, Berlin, 1981, pp. 217–246.

[3] D'Aristotile, A., Diaconis, P., and Freedman, D. *On merging of probabilities*. Sankhyā Ser. A **50**, 3 (1988), 363–380.

[4] Davydov, Y., and Rotar, V. *On asymptotic proximity of distributions*. J. Theoret. Probab. **22**, 1 (2009), 82–99.

[5] Dudley, R. M. *Convergence of Baire measures*. Studia Math. **27** (1966), 251–268. Correction: Studia Math. 51 (1974) 275.

[6] Dudley, R. M. *Distances of probability measures and random variables*. Ann. Math. Statist. **39** (1968), 1563–1572.

[7] Dudley, R. M. *Probabilities and metrics*. Matematisk Institut, Aarhus Universitet, Aarhus, 1976. Lecture Notes Series No. 45.

[8] Dudley, R. M. *Real analysis and probability*, Vol. 74. Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.

[9] Fabian, M., Habala, P., Hájek, P., Montesinos Santalucía, V., Pelant, J., and Zizler, V. *Functional analysis and infinite-dimensional geometry*. Springer-Verlag, New York, 2001.

[10] Gibbs, A. L., and Su, E. S. *On choosing and bounding probability metrics*. Internat. Statist. Rev. **70**, 3 (2002), 419–435.

[11] Gillman, L., and Jerison, M. *Rings of continuous functions*. Springer-Verlag, New York, 1976. Reprint of the 1960 edition, Graduate Texts in Mathematics, No. 43.

[12] Isbell, J. R. *Uniform spaces*. Mathematical Surveys, No. 12. American Mathematical Society, Providence, R.I., 1964.

[13] LeCam, L. *Convergence in distribution of stochastic processes*. Univ. Calif. Publ. Statist. **2** (1957), 207–236.

[14] LeCam, L. *Note on a certain class of measures*. Unpublished manuscript, 1970. http://www.stat.berkeley.edu/users/rice/LeCam/papers/classmeasures.pdf

[15] Pachl, J. *Measures as functionals on uniformly continuous functions*. Pacific J. Math. **82**, 2 (1979), 515–521.

[16] Rachev, S. T. *Probability metrics and the stability of stochastic models*. John Wiley & Sons Ltd., Chichester, 1991.

[17] Schaefer, H. H. *Topological vector spaces*. Springer-Verlag, New York, 1971. Third printing corrected, Graduate Texts in Mathematics, Vol. 3.

[18] van Handel, R. *Uniform observability of hidden Markov models and filter stability for unstable signals*. Ann. Appl. Probab. **19**, 3 (2009), 1172–1199.

[19] Zolotarev, V. M. *Modern theory of summation of random variables*. VSP, Utrecht, 1997.